

EXPLORATORY DYNAMIC CAPACITY ANALYSIS OF DEFENSE FORCES

Fred Ma

Centre for Operational Research & Analysis
Defence Research and Development Canada
National Defence Headquarters Carling
60 Moodie Drive, Ottawa, ON, Canada
Fred.Ma@forces.gc.ca

ABSTRACT

This paper describes DyFMCapA, a tool for analyzing a defense force's capacity to meet time-varying mission demands, and illustrates the exploration of metrics with which to inform the modification of a force's mix of assets. The importance of key requirements for agility in modeling and analysis is highlighted: a common modeling, simulation, data analysis, and visualization environment that supports a combination of relational data operations with array and imperative computation, and having mature programmatic and data interfaces to the prevailing data source environment.

KEYWORDS: force structure, force mix, force planning, capability based planning, capacity analysis

1 INTRODUCTION

Capability Based Planning (CBP) (Chief of Force Development 2014) refers to the process of systematically determining the most appropriate force options to meet national aims (Defense Terminology Database, Canada). In Canada's Department of National Defence (DND), the CBP process is overseen by Chief of Force Development, and has been supported by Centre for Operational Research & Analysis (CORA) through a number of iterations (Christopher et al. 2009; Rempel 2010; Pelletier 2013; Taylor 2017; Ma 2018b). Figure 1 shows the high level process for the current iteration. Step 1 produces departmentally approved *scenarios* [world situations that call for operational deployment by the Canadian Armed Forces (CAF)]. Step 2 determines possible CAF responses and estimates the assets needed (dubbed force elements, or *FElms*). The set of FElms needed for a given response to a scenario is dubbed a *force package (FP)*. Step 3 is force mix (FM) capacity analysis, which characterizes how well the CAF's FM (i.e., its inventory of FElms) is able to meet the FElm requirements of deployment to multiple concurrent operations. Step 4a identifies the need for various capabilities by the various scenario responses, and a subset of the response/capability combinations to be characterized in 4b. Step 4b examines the effectiveness with which an FP can deliver various capabilities needed by the corresponding scenario response, without considering capacity requirements. Step 5 combines the capacity and efficacy metrics from steps 3 and 4b to generate indicators of capability health.

This paper focuses on step 3. We describe a tool, coupled with split-apply-combine (SAC) and relational data paradigms, to analyze FElm resourcing of deployments when various scenarios randomly begin and end

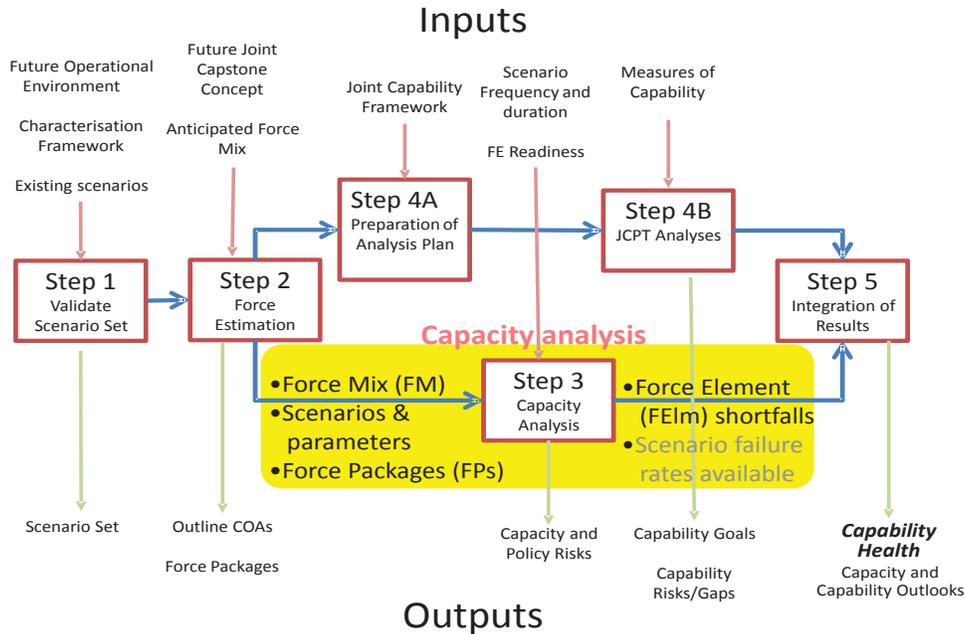


Figure 1: Initial phases of the CBP analysis (Taylor 2017). (JCPT = Joint Capability Planning Team)

along a timeline. The tool is dubbed DyFMCapA for Dynamic Force Mix Capacity Analyzer. We describe a light weight simulator at its core, a wrapper that orchestrates campaigns of simulations with which to explore what-if's, and demonstrate the use of the tool for exploratory analysis.

2 BACKGROUND

Two complementary types of FM capacity analysis are *static* and (ii) *dynamic* (Taylor 2008). Canada and allied nations have used both, although they are not as institutionalized in Canada (Pelletier 2009; Ma and Wesolkowski 2016; Ma and Dixon 2017) as they are in the US or the UK. *Static* capacity analysis provides indications of how well a FM is able to meet the FELm demands of *specifically identified combinations* of concurrent CAF operations, e.g., combinations deemed most likely or critical. *Dynamic* capacity analysis looks at how well a FM is able to meet overall demands for the various FELms as the demands vary with time – as various missions begin and end with unpredictable timing. It provides indicators of the capacity to meet the fluctuating demands that are likely to occur. DyFMCapA addresses *dynamic* capacity analysis.

The decision to develop DyFMCapA was informed by a study comparing three tools for dynamic analysis (Ma and Dixon 2017): Force Structure Readiness Assessment Tool (FSRA), Tyche, and Estimation of a Lower Bound for a Force Mix (ELBoFM). **FSRA** was developed through contracting to compare the supply and demand of FELms over time (de Souza 2015). **Tyche** was developed in CORA to simulate mission deployments at a high level of detail; with contractor support, it has evolved in sophistication and maturity over the years. **ELBoFM** is based on past models for estimating lower bound asset requirements when scheduling operations over a year. **All three** tools posed challenges for use in CBP. ELBoFM's lower bound model is overly optimistic when simulating operations over a period that is much longer than a year. Long simulation periods are needed, however, for characterizing steady-state supply and demand. The concerns about FSRA pertained to model transparency and tool operation, while the concerns for Tyche were about the requirements for high fidelity (mainly the level of detail of data).

The solution pursued was in-house development of the core simulator in DyFMCapA. It would have FSRA's fidelity and data requirements, and it would have the assured insight into the model and the robust operation

of Tyche. We would be free to: (i) choose an environment, and access all the internal data, for agile exploratory analysis; (ii) choose the desired output metrics; (iii) make ad hoc changes to DyFMCapA based on exploratory trials and a complete understanding of the model and implementation; and (iv) evolve DyFMCapA for future requirements, e.g., perhaps for algorithmic optimization of the FM.

3 CHOICE OF TOOL ENVIRONMENT

From past experience designing and characterizing dynamic FM capacity tools (Ma, Wesolkowski, and Eisler 2017, Ma and Dixson 2017), and modeling and analysis in general, we considered certain features of an analysis tool environment to be important to agile exploratory analysis. It should be geared to modeling, data analysis, and visualization via high level language constructs. This avoids high volume recoding and verification when evolving the modeling and analysis in way that is conceptually simple. More sophisticated environments require greater expertise and familiarity, however, so it makes less sense to straddle multiple environments, e.g., as one might for a third generation language. That is, it can be advantageous to leverage and expand expertise in a particular environment, consolidating as much analysis therein as practical. This minimizes time spent defining and/or implementing interfaces and/or adapters for them, and the attendant lost in agility. It is important, however, to have well developed interfaces to the environments where data is from, and to where data will go.

Microsoft Office tools are pervasive in DND, so the chosen environment had to have good interfaces to Excel and Access. CORA is part of Defence Research and Development Canada (DRDC), where Matlab is pervasive; it was a natural choice for the computational environment. There are simple functions to import and export data to/from Excel. The requirements are more complex, however, for campaigns of simulations that sweep through suites of manually prepared data. The Component Object Model (COM) interface (The Mathworks 2019) was critical in alleviating the manual, error-prone tracking of spreadsheets and bounding coordinates of many heterogeneous data sets, especially when the data sets are subject to changes in size. COM lets Matlab control Excel with flexibility comparable to Visual Basic for Applications; this reduces bookkeeping by enabling the programmatic access to List Objects by name [a List Objects is Excel's high level construct for intelligently manipulating data sets (Newman 2014)].

While Excel offered an extremely flexible data source/sink environment, the SAC computing paradigm (Wickham 2011), which is routine in relational database (RDB) analysis, was essential to DyFMCapA's internal analysis agility, as were relational operations in general. SAC code patterns are powerful for data mining; they make it easy to reduce the dimensionality of data by applying arbitrary aggregation functions over a subset of the dimensions. This avoids lengthy one-off coding for exploring different aggregations of the FEIm assignment data from simulating the resourcing of deployments. The initial plan for getting RDB functionality was through Access, to which Structured Query Language (SQL) code is conveyed as strings. Those who have done this, however, would probably agree that this impedes agility for complex, nested queries. Fortunately, we found Matlab's native relational data operations to be fairly complete, e.g., *findgroups*, *splitapply*, *accumarray*, *rowfun*, *unstack*, variations on *join*. While not as performance-optimized as SQL engines for SAC operations with a lot of splitting, we did not have a situation that warranted the code and agility overhead of resorting to SQL. Rather, keeping everything native allowed for the powerful blending of imperative and array computation with RDB operations, often within the same expressions. It also avoided the need for complex database (DB) interface parameters that depend on the operating system and the DB; this made the implementation less fragile.

4 CORE SIMULATOR

Model: DyFMCapA simulates a collection of scenarios, each of which recur with random timing over a long period of time. Depending on a scenario's Poisson rate and duration distribution, it is not uncommon

for its occurrences to span multiple time steps. In each time step, DyFMCapA assigns FELms to fulfill the FP requirements for scenarios that are occurring in that time step. The latter creates the demand for FELms while the FM serves as the supply. Hence, the FM must specify the FELms that a nation's defense forces possess, and their deployable quantities. The latter limits the degree to which the demand can be satisfied within each time step. The FM can designate some FELms as always available, e.g., teams/headquarters assembled ad hoc, as needed, and services supplied to missions from outside of theater.

On the demand side, a set of scenarios needs to be defined, representing various types of operations to which the forces deploys, e.g., peace missions, Humanitarian Operations & Disaster Relief, combat, etc. The data required for each scenario includes: (i) its Poisson rate; (ii) a triangular or flat probability distribution for the duration of the scenario; (iii) its *FP*, which specifies the FELms needed to deploy to the scenario, and their quantities; and (iv) the scenario's priority, which determines the order in which scenario demands are satisfied within each time step. Since FELms deployed to one mission could be re-deployed to different missions in successive time steps, the time step should be chosen to be long enough for this to be realistic. Each scenario can optionally be designated as "baseline"; these are treated as always occurring rather than as a Poisson process.

Implementation: Taking advantage of SAC processing requires that the FELm assignments be captured in an RDB-style data cube. Table 1 shows a sample of records from a very large table. For each time step value, there are different rows for the different scenarios occurring in that time step, identified by integers. For each unique combination of (*Time step, Scenario*), there may be different rows for different occurrences of that scenario, also identified by integers. There may be different occurrences of a single scenario within a time step is because: (i) a Poisson process can have multiple events in the finite duration of a time step and (ii) quite apart from the finite time step duration, a Poisson process represents randomly timed events with infinitesimal time durations, whereas our scenarios have finite durations. With regard to (ii), an occurrence of scenario *x* might not have ended before the next begins. For each unique combination of (*Time step, Scenario, Occurrence*), there are multiple rows, one for each FELm that needs to be sent on that mission, i.e., one row for each FELm in the scenario's FP. Each such row also contains the quantity required (*Demand*), as per the FP, and the quantity that DyFMCapA was able to assign (*Supply*).

Each of the first four columns, (*Time step, Scenario, Occurrence, and FELm*) can be considered an axis in a 4D data cube. Our aim was to leverage the simplicity of SAC for aggregations based on unique values of any of these fields, or unique combinations thereof.

5 CAMPAIGN ORCHESTRATION WRAPPER

For a one-off simulation, a lot of time can be spent setting up data, collating output data, and converting it into information. Exploring trends and what-if's, however, can entail campaigns of simulations for multidimensional sweeps through alternative data sets and/or stepwise perturbations thereof. The need for manual manipulation is much less tolerable due to the compounding of time needed, tedium, and likelihood of error. Campaigns of simulations are also important for confirming sensible behavior for a new model, implementation, or modifications thereof. The troubleshooting of errors arising from poor automation support for simulation campaigns can significantly increase the time required – and that is for the fortunate cases in which erroneous results are recognized.

DyFMCapA mitigates this by using a simulation control table, which acts as the control console for the campaign. In it, an analyst lists the data sets to simulate by name. There can be multiple data sets for FMs, FPs, and scenario timing parameters (Poisson rates, duration distributions), i.e. a 3D cube of data sets. The FM data sets represent different FMs that a nation could plan for. The different FP data sets could account for uncertainty in the scenario requirements, or alternative world situations in which scenario requirements are different. The different sets of scenario timing parameters are meant to reflect alternative futures in

Table 1: Example entries from data cube for assignment of FEIms to scenarios.

Time step	Scenario	Occurrence	FEIm	Demand	Supply
1	1	5	9	5	5
1	6	3	21	3	3
3	6	3	23	1	0
8	6	4	13	1	0
9	4	11	26	1	1
9	6	4	21	3	3
10	5	2	28	3	3

which (say): (i) combat missions are more frequent, and peace missions less so, or (ii) the converse. The input data environment is Excel, which makes it very easy to browse the data sets and format them in any way that improves the sense-making.

In capacity analysis, the main decision variables are the quantity of each FEIm in the FM. These variables are the ones most obviously influenced through planning and resourcing. For focused sweeps of variations on each FM data set, one may additionally specify a FM perturbation table. It identifies a subset of the FEIms in the FM to perturb, and lists how their FM quantities are perturbed in each perturbation step. Hence, the cube of data sets is actually 4D. This provides analytical flexibility and a common output data format regardless of how the dimensions are swept, thus simplifying the development of post-processing. To avoid combinatorial explosion, however variations along each of the four dimensions need to be chosen judiciously.

Post-simulation modules generate and compare plots of the output metrics (Section 6) for the simulations in the campaign. The simulation control table also specifies the data sets for which the output metrics are generated for comparison. Each module extracts the right simulation results with which to generate the output metrics. These results consist mostly of partial aggregations of FEIm assignment data (Table 1), e.g., aggregated by FEIm, scenario, unique combinations of FEIm and time step, etc. When a module extracts the specified subset of partial aggregations, it avoids navigating multiple data files and delving into nested arrays of objects; instead, we maintain the simplicity of a single output data cube for each of the partial aggregations by adding four extra fields to identify which data set in the 4D input cube the results are for. Set up in this way, extracting the right subset of rows is trivial in most 4G analysis environments.

6 OUTPUT METRICS

The definitions of the output metrics were inspired by the those from FSRA and Tyche. Those tools generate complementary indicators of capacity falling into two categories: (i) per-FEIm statistics and per-scenario statistics. In DyFMCapA, the different metrics and comparisons thereof are generated by different post-simulation modules.

FEIm satisfaction time (FST) is the *percentage* of the *time* steps for which the FM’s supply of a given FEIm meets the overall demand for that FEIm. This metric has *optimistic* and *pessimistic* versions. Optimistic FST (oFST) includes all time steps for which there is zero demand for the FEIm in question, and counts these as time steps for which demand is satisfied. Pessimistic FST (pFST) only looks at time steps for which there is non-zero demand for a given FEIm. Both are informative; oFST gives an indication of the proportion of all time in which the demand for an FEIm goes unmet, while pFST indicates how often demand will go unmet *when the FEIm is needed*. Unless otherwise noted, the latter is the default.

FEIm demand relative to supply, or just **relative demand**, indicates the degree to which the quantity of a FEIm in the FM (i.e., the supply) either falls short of, or exceeds, overall demand. Within each time step,

for each FEIm, the ratio of total demand over total supply is generated. For each FEIm, this data then is then averaged over all time steps in which there is demand for the FEIm.

Scenario fulfillment is a per-scenario metric representing the proportion of a scenario’s FEIm requirements that are fulfilled, averaged over all FEIm in the scenario’s FP and all time steps of all scenario occurrences. For a given scenario s , the average taken is the median (Graham, Knuth, and Patashnik 1994) of ratio *Supply/Demand* generated from all records in Table 1 that specify s in the *Scenario* field. We chose the median based on intuitive suitability; the end effect is simply the ratio of all step-wise unit FEIm supplied to all step-wise unit FEIm demanded. Not all FEIm are equivalent in importance to each scenario, however, so the FP data set includes a field to specify scenario-specific FEIm weights. For illustrative purposes, these weights are given default values equivalent to the inverse of the quantity of the FEIm in the FM, $n_{\text{FEIm}}^{\text{FM}}$. If the FP requires $n_{\text{FEIm}}^{\text{FP}}$ of a given FEIm, therefore, the weight is $n_{\text{FEIm}}^{\text{FP}}/n_{\text{FEIm}}^{\text{FM}}$, which can be seen as the fraction of the "fleet" required ("fleet" being interpreted loosely, as it can include things like headquarters, of which there may be only one). The metric $1/n_{\text{FEIm}}^{\text{FM}}$ is an indicator of the FEIm’s *scarcity*. We may view $n_{\text{FEIm}}^{\text{FP}}/n_{\text{FEIm}}^{\text{FM}}$ as the FEIm’s contribution to the difficulty of generating the FP, i.e., an FP-specific *FEIm difficulty*.

7 APPLICATION OF DYFMCAPA

7.1 Data sets

Data sets for CBP are nontrivial to develop. They are generated through a consultative process involving high level parts of DND. DyFMCAPA was applied to two data sets: (i) 2015-era data (Ma and Dixon 2017), consisting of approximately 80 FEIm and 8 scenarios, and (ii) a nascent higher-resolution 2018 data set meant for the current CBP cycle, consisting of approximately 385 FEIm and 60 scenarios. Having been used in past studies, the 2015 data is more mature, containing 3 FMs and 3 sets of stochastic timing parameters per scenario (Eisler 2015; Ma and Wesolkowski 2016; Ma and Dixon 2017). It also contains up to 2 contingent FPs per scenario for use in Tyche’s more detailed modeling.

We wanted to explore analyses that could be informative not just for the two data sets above, but for future data sets. However, these two data sets have very specific patterns of scenario difficulties and priorities. Therefore, the data sets were varied in ways that attempted to preserve their realistic characteristics. We will describe analyses of an illustrative subset of the data set variations .

7.2 Comparative plots

This section shows examples of comparisons of the output metrics, which have been preliminarily reported internally (Ma 2018a). The results are from the 2015 data, as they have the diversity with which to compare variations in FMs, FPs, and scenario timing parameters.

Figure 2 shows relative demand of FEIm, which reflects their shortfall or excess. A pair of bars is shown for each FEIm. The blue bar shows the average demand for that FEIm *relative* to the quantity in the As-Is FM. These As-Is quantities prefix the FEIm names along the x -axis. Bars that extend above the 100% line identify FEIm whose FM quantities fall short of demand, on average. The yellow bars show FEIm demand relative to a better resourced “warfighting” FM, wherein FEIm whose names are suffixed by triple-asterisks have the increased quantities (shown after the asterisks). As expected, the relative demands fall in proportion to the increased quantities in the FM, as do any deficits.

Figure 3 shows FST time as a function of two FMs. Recall that this is the percentage of time that the FM is able to satisfy overall demand for each FEIm, i.e., for which there is no “overage” (as annotated in the figure). There are twice as many bars as in Figure 2 because this metric has optimistic and pessimistic

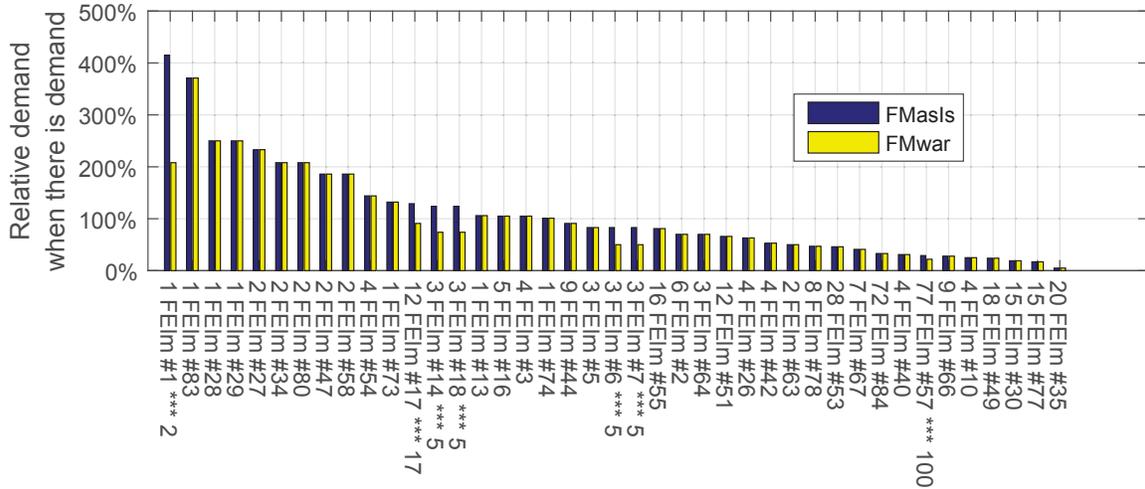


Figure 2: Relative demand as a function of two FMs.

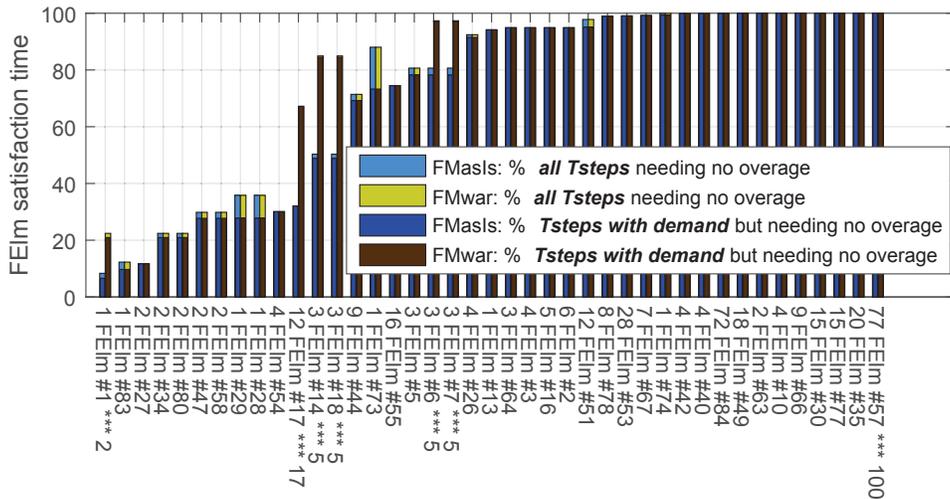


Figure 3: Pessimistic and optimistic FST time as a function of two FMs.

versions. The pessimistic bars are shorter, so they are the darker set of bars in the foreground, while the lighter optimistic pairs of bars are directly behind them. As was the case in Figure 2, the left bars for each FEIm are for the As-Is FM while the right bars are for the Warfight FM. Not surprisingly, the latter show better FST for FEIm whose quantities in the FM are increased.

Figure 4 (left) shows the result of incrementing selected FEIm quantities in the FM. As expected, the pFST increases. Each curve corresponds to one of the selected FEIm, and the increasing quantities are shown on the curves themselves, adjacent to each data point. FEIm #57 does not show any improvement because its demand are already satisfied 100% of the time, even without incrementation.

Figure 5 shows the sensitivity of pFST to changing FP composition. The FPs for the three most demanding scenarios (5a, 5b, and 5c) are perturbed by halving and doubling all the required FEIm quantities. Some of the FEIm are not affected by the perturbation because not all FEIm appear in the perturbed FPs.

Figure 6 shows pFST for three alternative futures: (i) a *peace* future in which peace missions are more prevalent and warfighting missions are less so; (ii) an *As-Is* future in which the various mission types have

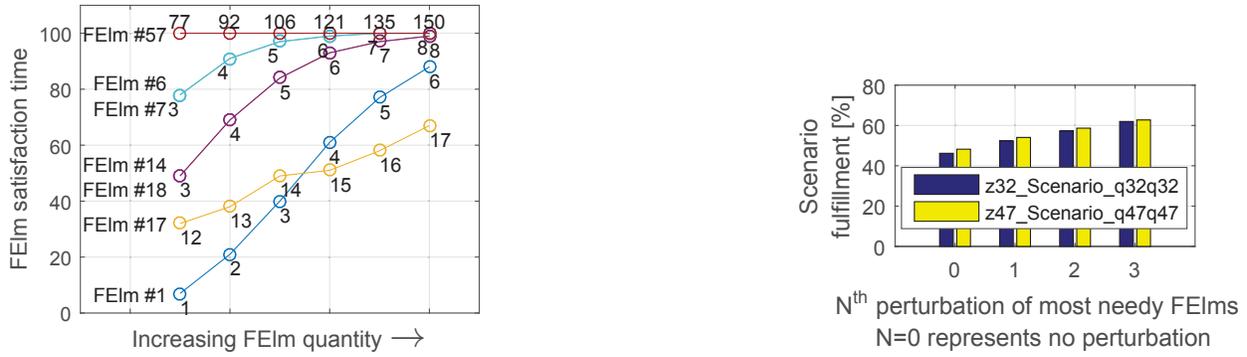


Figure 4: Increasing FELm quantities within the FM. Left: 2015 data. Right: 2018 data.

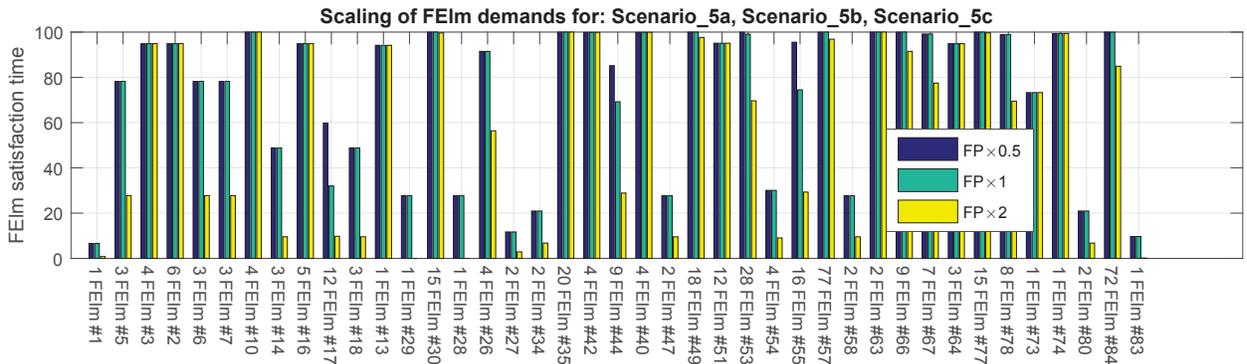


Figure 5: Sensitivity of pFST to FP composition.

the same prevalence as they have had in the past; and (iii) a *war* future in which warfighting missions are more prevalent, and peace missions are less so. Note that “peace” missions (e.g., peace enforcement, peace keeping) require the readiness to use force, and hence, require FELms normally thought of as combat related. As shown in Figure 6, prevalence of peace missions can actually deplete the FM, as evidenced by the lower pFST (the blue bars are near zero for many FELms).

Figure 7 (left) shows scenario fulfillment for the 8 scenarios in the 2015 data. To find a pattern between the fulfillments and possible causal factors, we placed the fulfillment bars on a 2D map of the scenarios’ priorities and their difficulties (the floor of the graph), where a scenario’s difficulty is simply the sum of the FELm difficulties for the FELms in the FP. Scenario difficulty was examined as a second causal factor because we did not see the expected increase of fulfillment with increasing scenario priority. Unfortunately,

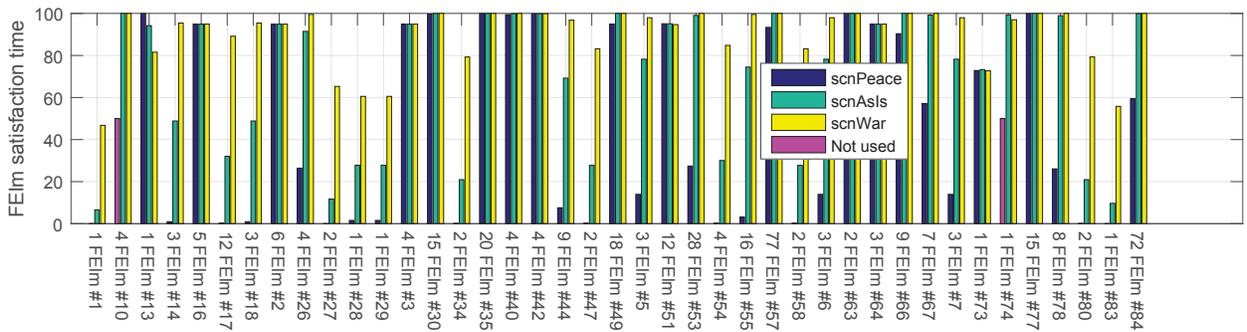


Figure 6: pFST for peaceful/war-like future.

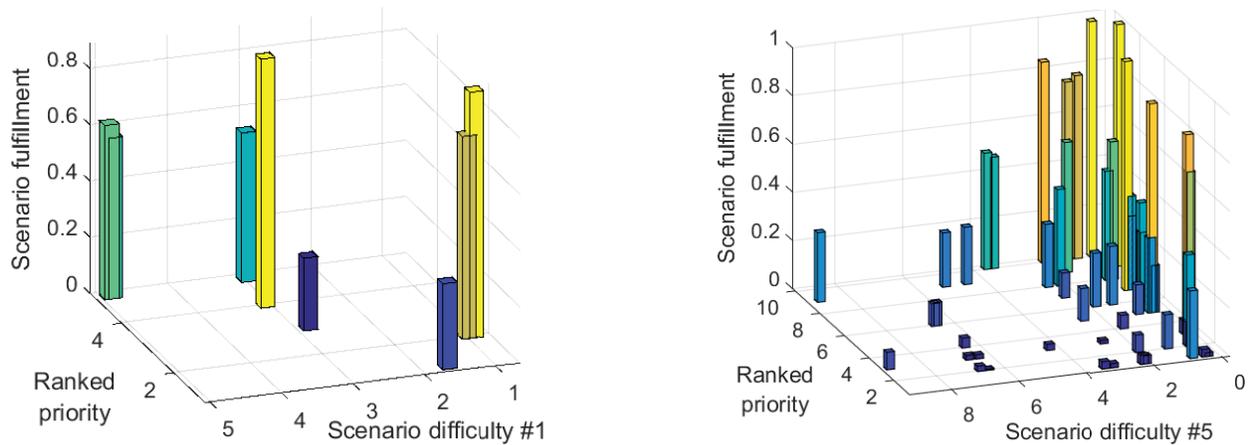


Figure 7: Scenario fulfillment. Left: 2015 data. Right: 2018 data.

the scenarios are not well positioned on the priority-difficulty plane to generate insight. We expect low scenario fulfillments at the front corner of the graph, where scenarios are most difficult and have the lowest priority; conversely, the highest fulfillments should be at the back corner. The actual scenarios are not positioned along this diagonal; they lie roughly along the perpendicular diagonal, where variation in priority and difficulty have opposing effects. Hence, we were still in need of a way to tie scenario fulfillment to FELm difficulty, as that would provide indications of how to shore up the FM based on which scenarios were considered most important to improve.

7.3 Ad hoc exploratory analysis

The importance of an agile analysis environment was highlighted by the ad hoc analysis to find a link between scenarios fulfillment and FELm difficulties. Due to its small number of scenarios, the 2015 data was not ideal even if the scenarios were well positioned in Figure 7. Hence, we focused on the 2018 data. However, all the scenarios therein were low in difficulty and/or low in priority. In order to explore possible ways to link fulfillments to FELms, we needed artificial data that was more diverse, yet retained some of the real data's characteristics. The simplest way to do this was randomize the priorities, thereby scattering the scenarios across the difficulty-priority plane. Any schemes that behaved sensibly for the 2018 data could be corroborated using artificial data generated from the 2015 data, e.g., the low number of 2015 scenarios could be amplified in a realistic way by turning each scenario's contingent FP into its own scenario, albeit with a reduced Poisson rate to preserve the overall number of occurrences.

Our exploration examined different ways of quantifying FELm difficulty and aggregating them into scenario difficulty. If the resulting pattern of scenario fulfillment bars was sensible, then the scheme was considered a good candidate, i.e., if there was a general trend of increasing fulfillment with increasing priority and decreasing scenario difficulty. The clearer this trend, the better. For scenarios of particular interest, the FELm difficulty metric could then be used to rank the FELms within the FPs for further attention, e.g., perhaps by simulating the effect of increasing FM quantities for difficult FELms. We found four good schemes for quantifying FELms difficulty and scenario difficulty.

The first two schemes are the simplest. Scheme **S1** was to sum up the FELm difficulties within a FP, as above. The FELm difficulty metric therein, however, only considered the fleet size in the FM, i.e., the *supply* side. Actual difficulty needs to also consider the competing *demand* for a FELm. DyFMCapA already generates a good metric for this: relative demand. Scheme **S2** uses this for FELm scarcity. For clarity, we continue

to refer to the metric for the FELm alone as *scarcity*, while *FELm difficulty* refers to the FP-specific metric that also factors in the quantity required by the FP, and *scenario difficulty* is the aggregation of the FELm difficulties over a FP.

Scheme **S3** was based on the reliability engineering metric for *availability* (Ebeling 1997): the percentage of the unit demands for a FELm that were met. This is easy to tally up from Table 1; the question was how to meaningfully aggregate the metrics over a FP. A pessimistic approach views the FELms as non-redundant systems, and multiplies together the metric $A^{n_{\text{FELm}}^{\text{FP}}}$ for each FELm in a FP, where A is the FELm's availability (common across FPs). The inverse of the product might then be a good scenario difficulty metric, which essentially means that the FELm difficulty is $1/A^{n_{\text{FELm}}^{\text{FP}}}$. This application of availability is not strictly correct; apart from the non-redundancy assumption, the chances of satisfying the $i + 1^{\text{st}}$ demand for a FELm in the FP is not independent of the chances of satisfying the i^{th} demand (independence is a prerequisite for multiplicative aggregation). The aim, however, was not to develop a reliability model; it was to find an empirically well behaved link between scenario fulfillment and FELm difficulty.

We found S3 to be so pessimistic that the multiplicative aggregation into scenario difficulty led to numerical underflow, especially for the larger FPs in the 2018 data. We avoided this by equivalently summing in the log domain; the FELm difficulty thus became $-n_{\text{FELm}}^{\text{FP}} \log A$. While the result was too negative to meaningfully transform back into the linear domain, it turned out to be a good scenario difficulty metric without the transformation. We treated the metric as interval data, so its negativity is not an issue.

Scheme **S4** was motivated by recognizing oFST as *marginal* availability. Its definition (page 5) implies that, for an arbitrarily chosen time step, oFST is the probability that the request for one more unit of a FELm will be met. The FELm scarcity is then *unavailability* $u = 1 - \text{oFST}$, the FELm difficulty is $n_{\text{FELm}}^{\text{FP}} \cdot u$, and the scenario difficulty is the sum of the latter over the FP.

Reconciling scenario difficulties and ranking the FELms: Subjectively, S1 through S4 (S1-4) seemed equally good. Figure 7 (right) shows an amalgamation that reflects the four schemes well. Each scenario's difficulty metric is the average across S1-4. Since we were seeking a scheme that produced a sensible pattern, we had the liberty of scaling the scenario difficulty values from each S1-4 scheme so that their ranges matched the priority axis in the graph. This helped in the inspection and sense-making, and it also ensured that the four sets of results had comparable effect on the average.

The FELms within each FP then needed to be ranked by FELm difficulty in order to focus the attention of decision makers (DMs). S1-4 yield four rankings, which we reconciled into a "consensus" ranking. Since the purpose was to bring attention to the most needed FELms rather than making a decision, we could use a simple consensus scheme based primarily on intuition: Borda-x, a modification of Borda counting (Ma et al. 2013). Within a given S1-4 scheme, if there was a tie between FELms that would otherwise occupy ranks p through q , they all take on the mean rank $(p + q)/2$. Each FELm is then given a score that is the average of its Borda-x values across S1-4. The consensus ranking is then determined based on these scores. For convenience, initially, we "normalized" the Borda-x values prior to averaging so that they occupied the range $[0,1]$ rather than $[1, n_{\text{FELm}}^{\text{FP}}]$. The average, which we refer to as the S1-4 score, tends to be in the middle of $[0,1]$ due to averaging.

As an example of how the ranking of FELms within FPs might inform what-if investigations involving multiple scenarios, consider a subset of scenarios that might be of particular interest, e.g., the two green bars in Figure 7 (right) with high priority=10 and difficulty ≈ 3.5 . The scenarios may have different FPs, with different consensus rankings of different FELms. This poses conceptual challenges in consolidating the rankings. Our approach was: for each distinct FELm, further average the S1-4 scores across all FPs that contain that FELm. Essentially, each FP that contains the FELm provides an "opinion" of where the FELm sits within "the pack", and multiple opinions are averaged. We did not re-rank the S1-4 scores before averaging

across scenarios because we wanted to retain the information behind their central tendency, i.e., either a consensus across S1-4 on a middle position, or a lack of consensus on non-middle values.

While the two green scenarios above have similar FPs, $n_{\text{FElm}}^{\text{FP}}$ can be very different across FPs in general. Care may be needed in the above normalization of Borda-x values to ensure sensibility in the subsequent averaging of S1-4 scores across scenarios. Consider two hypothetical scenarios, sc4 and sc99, with 4 and 99 FElms in their respective FPs. For sc4's FP, we want the FElm ranked at (say) 2 to normalize to $(0.25 + 0.50)/2$, i.e., the middle of the second quartile. Otherwise, it's position will be biased relative to the sc99 FElms in the second quartile. For an arbitrary FElm with rank $r \in [1, n_{\text{FElm}}^{\text{FP}}]$, the normalization $(r - 0.5)/n_{\text{FElm}}^{\text{FP}}$ has the desired behavior.

For the two green scenarios mentioned, the 20 most difficult FElms had quantities of 1 to 3 in the FM. We investigated the effect of boosting these quantities. We used DyFMCapA's FM perturbation table (Section 5) to specify 4 perturbation steps for these FElms, incrementing their quantities by 1 in each step. Figure 4 (right) shows the increase of fulfillment with the increasing FElm quantities. In practice, we expect the DM(s) and supporting experts to help identify practical combinations of FElm increases for investigation.

8 CONCLUSIONS AND FUTURE WORK

Our dynamic FM analysis environment currently has a light weight simulator at its core, which is used within a simulation campaign wrapper. A key objective was agile exploration of different data sets across different parameter dimensions, minimizing the burden of managing such multi-dimensional arrays of data sets at the input and output, and minimizing time consuming errors. Another objective was agility in exploratory analysis. Both objectives were enabled by designing to exploit SAC and relational data operations. An underlying environment was chosen to provide a powerful combination of the above with array and imperative computation. This removed the need to interface across multiple environments, speeding up the translation of ideas to implementation. We presented an assortment of comparisons of outputs from campaign sweeps, and described the exploration of metrics that provide indications of what FM modifications could improve the fulfillment of FElm demands of selected scenarios.

Potential improvements being considered are: (i) modeling scenario setup and close-out phases so that FElms specialized for those activities aren't tied up for the entire scenario; (ii) investigating whether Pareto front concepts can provide better indications of which FElms to increase the quantity of; (iii) heuristic search for a good FM, along with sensible objectives and constraints.

ACKNOWLEDGMENTS

Ben Taylor helped contextualize the dynamic FM capacity problem and create opportunities for this work. Murray Dixon helped frame the features for tool comparisons, and diversify the 2015 data set. Cheryl Eisler is our Tyche expert and was key in creating the original 2015 data. We are indebted to Directorate of Capability Analysis for the 2018 data.

REFERENCES

- Chief of Force Development 2014, June. *Capability Based Planning Handbook*.
- Christopher, G., D. Blakeney, R. Petryk, B. Taylor, L. Kerzner, A. Beard, V. Fong, and M. Ball. 2009, December. "Strategic Capability Roadmap Version 1.0: Analytic Framework". Technical report, CORA, DRDC, Canada.
- de Souza, S. 2015, August. "User Guide: Force Structure Readiness Assessment Tool". Contractor report, DRDC CORA, Canada.

- Ebeling, C. E. 1997. *An introduction to reliability and maintainability engineering*. Long Grove IL, Waveland Press, Inc.
- Eisler, C. 2015, March. “Force Structure Evaluation for Capability-Based Planning”. Scientific report, DRDC CORA, Canada.
- Graham, R. L., D. E. Knuth, and O. Patashnik. 1994. *Concrete Mathematics: A Foundation for Computer Science*. 2 ed. Addison-Wesley.
- Ma, F. 2018a, September. “Dynamic Force Mix Capacity Analyzer (DyFMCapA): Toward Capability Based Planning (CBP) Step 3”. Scientific letter, DRDC CORA, Canada.
- Ma, F. 2018b, April. “Force mix analysis and war gaming to support Capability Based Planning: Review of methods and implications for way ahead”. Reference document, DRDC CORA, Canada.
- Ma, F., M. G. Ball, E. Pelletier, and M. Lessard. 2013, May. “Examination of the Prioritization Problem in the Context of C4ISR Project Prioritization for CFD”. Technical memorandum, DRDC CORA, Canada.
- Ma, F., and M. Dixon. 2017, June. “A comparison of strategic force mix tools for application to capability based planning”. Scientific report, DRDC CORA, Canada.
- Ma, F., and S. Wesolkowski. 2016. “Estimating Force Mix Lower Bounds Using a Multi-objective Evolutionary Algorithm”. In *IEEE Symposium Series on Computational Intelligence*. Athens, Greece, IEEE, IEEE.
- Ma, F., S. Wesolkowski, and C. Eisler. 2017, March. “ELBoFM: Estimation of a Lower Bound for a Force Mix: An evolutionary multi-objective trade-off model”. Scientific report, DRDC CORA, Canada.
- Newman, C. 2014, June. “The VBA Guide To ListObject Excel Tables”. <http://www.thespreadsheetguru.com/blog/2014/6/20/the-vba-guide-to-listobject-excel-tables>, accessed 2019-01-25.
- Pelletier, E. 2009, October. “The Concurrency Analysis Tool: A Tool from the Strategic Capability Roadmap version 1.2 Toolsuite”. Technical memorandum, DRDC CORA, Canada.
- Pelletier, E. F. 2013, October. “New Methods and Tools for Capability Based Planning: Tools used in Spirals 1 and 2 of Capability Based Planning 2012”. Technical memorandum, DRDC CORA, Canada.
- Rempel, M. 2010, September. “An Overview of the Canadian Forces’ Second Generation Capability-Based Planning Analytical Process”. Technical memorandum, DRDC CORA, Canada.
- Taylor, B. 2008. “TTCP JSA TP-3 Workshop - Force Structure Analysis”. Technical report, The Technical Cooperation Program.
- Taylor, B. 2017, August. “Toward an Enhanced Capability Based Planning Approach”. Reference document, DRDC CORA, Canada.
- The Mathworks 2019. “Call COM Objects”. <http://www.mathworks.com/help/matlab/using-com-objects-in-matlab.html>, accessed 2019-01-25.
- Wickham, H. 2011, April. “The Split-Apply-Combine Strategy for Data Analysis”. *Journal of Statistical Software* vol. 40 (1). <http://www.jstatsoft.org/article/view/v040i01/v40i01.pdf>.

AUTHOR BIOGRAPHIES

FRED MA is a defense scientist at DRDC CORA. He holds a PhD in Electrical Engineering. His research interests lie in applied operational research. His email address is Fred.Ma@forces.gc.ca.